

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT:
 - (A) NAME: Middeldorp, Jaap Michiel.
- (ii) TITLE OF INVENTION: Peptides and nucleic acid sequences related to the Epstein-Barr virus.
- (iii) NUMBER OF SEQUENCES: 22
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Organon Teknika Corporation
Biotechnology Research Institute
 - (B) STREET: 1330-A Piccard Drive
 - (C) CITY: Rockville
 - (D) STATE: Maryland
 - (E) COUNTRY: USA
 - (F) ZIP: 20850-4377
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: Patentin Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: EP 92200721.6
 - (B) FILING DATE: 13-MAR-1992
 - (C) CLASSIFICATION
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Bobrowicz, Donna
 - (B) REGISTRATION NUMBER: 32,196

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 538 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Epstein-Barr virus
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CATGATGGCA CGCCGGCTGC CCAAGCCCAC CCTCCAGGGG AGGCTGGAGG CGGATTTTCC 60
 AGACAGTCCC CTGCTTCCTA AATTTCAAGA GCTGAACCAG AATAATCTCC CCAATGATGT 120
 TTTTCGGGAG GCTCAAAGAA GTTACCTGGT ATTTCTGACA TCCCAGTTCT GCTACGAAGA 180
 GTACGTGCAG AGGACTTTTG GGGTGCCTCG GCGCCAACGC GCCATAGACA AGAGGCAGAG 240
 AGCCAGTGTG GCTGGGGCTG GTGCTCATGC ACACCTTGGC GGGTCATCCG CCACCCCCGT 300
 CCAGCAGGCT CAGGCCGCCG CATCCGCTGG GACCGGGGCC TTGGCATCAT CAGCGCCGTC 360
 CACGGCCGTA GCCCAGTCCG CGACCCCCTC TGTTTCTTCA TCTATTAGCA GCCTCCGGGC 420
 CGCGACTTCG GGGGCGACTG CCGCCGCCTC CGCCGCCGCA GCCGTCGATA CCGGGTCAGG 480
 TGGCGGGGGA CAACCCCACG ACACCGCCCC ACGCGGGGCA CGTAAGAAAC AGTAGCCC 538

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 176 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met	Ala	Arg	Arg	Leu	Pro	Lys	Pro	Thr	Leu	Gln	Gly	Arg	Leu	Glu	Ala	1	5	10	15
Asp	Phe	Pro	Asp	Ser	Pro	Leu	Leu	Pro	Lys	Phe	Gln	Glu	Leu	Asn	Gln	20	25	30	
Asn	Asn	Leu	Pro	Asn	Asp	Val	Phe	Arg	Glu	Ala	Gln	Arg	Ser	Tyr	Leu	35	40	45	
Val	Phe	Leu	Thr	Ser	Gln	Phe	Cys	Tyr	Glu	Glu	Tyr	Val	Gln	Arg	Thr	50	55	60	
Phe	Gly	Val	Pro	Arg	Arg	Gln	Arg	Ala	Ile	Asp	Lys	Arg	Gln	Arg	Ala	65	70	75	80
Ser	Val	Ala	Gly	Ala	Gly	Ala	His	Ala	His	Leu	Gly	Gly	Ser	Ser	Ala	85	90	95	
Thr	Pro	Val	Gln	Gln	Ala	Gln	Ala	Ala	Ala	Ser	Ala	Gly	Thr	Gly	Ala	100	105	110	

Leu Ala Ser Ser Ala Pro Ser Thr Ala Val Ala Gln Ser Ala Thr Pro
 115 120 125
 Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala
 130 135 140
 Thr Ala Ala Ala Ser Ala Ala Ala Val Asp Thr Gly Ser Gly Gly
 145 150 155 160
 Gly Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln
 165 170 175

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1038 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGCTATCAG GTAACGCAGG AGAAGGAGCA ACAGCCTGCG GAGGTTCGGC CGCCGCGGGC 60
 CAGGACCTCA TCAGCGTCCC CCGCAACACC TTTATGACAC TGCTTCAGAC CAACCTGGAC 120
 AACAAACCGC CGAGGCAGAC CCCGCTACCC TACGCGGCCC CGCTGCCCCC CTTTTCCCAC 180
 CAGGCAATAG CCACCGCGCC TTCCTACGGT CCTGGGGCCG GAGCGGTTCGC CCCGGCCGGC 240
 GGCTACTTTA CCTCCCCAGG AGGTTACTAC GCCGGGCCCC CGGGCGGGGA CCCGGGTGCC 300
 TTCTTGCGGA TGGACGCTCA CACCTACCAC CCCACCCAC ACCCCCCTCC GGCCTACTTT 360
 GGCTTGCCGG GCCTCTTTGG CCCCCCTCCA CCCGTGCCTC CTTACTACGG ATCCCACTTG 420
 CGGGCAGACT ACGTCCCCGC TCCCTCGCGA TCCAACAAGC GGAAAAGAGA CCCCAGAGGAG 480
 GATGAAGAAG GCGGGGGGCT ATTCCCGGGG GAGGACGCCA CCCTCTACCG CAAGGACATA 540
 GCGGGCCTCT CCAAGAGTGT GAATGAGTTA CAGCACACGC TACAGGCCCT GCGCCGGGAG 600
 ACGCTGTCCT ACGGCCACAC CGGAGTCGGA TACTGCCCCC AGCAGGGCCC CTGCTACACC 660
 CACTCGGGGC CTTACGGATT TCAGCCTCAT CAAAGCTACG AAGTGCCCAG ATACGTCCCT 720
 CATCCGCCCC CACCACCAAC TTCTCACCAG GCAGCTCAGG CGCAGCCTCC ACCCCCGGGC 780
 ACACAGGCCC CCGAAGCCCA CTGTGTGGCC GAGTCCACGA TCCCTGAGGC GGGAGCAGCC 840

GGGAACTCTG GACCCCGGGA GGACACCAAC CCTCAGCAGC CCACCACCGA GGGCCACCAC 900
 CGCGGAAAGA AACTGGTGCA GGCCTCTGCG TCCGGAGTGG CTCAGTCTAA GGAGCCCACC 960
 ACCCCCAAGG CCAAGTCTGT GTCAGCCCAC CTCAAGTCCA TCTTTTGCGA GGAATTGCTG 1020
 AATAAACGCG TGGCTTGA 1038

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Leu	Ser	Gly	Asn	Ala	Gly	Glu	Gly	Ala	Thr	Ala	Cys	Gly	Gly	Ser	1	5	10	15
Ala	Ala	Ala	Gly	Gln	Asp	Leu	Ile	Ser	Val	Pro	Arg	Asn	Thr	Phe	Met	20	25	30	
Thr	Leu	Leu	Gln	Thr	Asn	Leu	Asp	Asn	Lys	Pro	Pro	Arg	Gln	Thr	Pro	35	40	45	
Leu	Pro	Tyr	Ala	Ala	Pro	Leu	Pro	Pro	Phe	Ser	His	Gln	Ala	Ile	Ala	50	55	60	
Thr	Ala	Pro	Ser	Tyr	Gly	Pro	Gly	Ala	Gly	Ala	Val	Ala	Pro	Ala	Gly	65	70	75	80
Gly	Tyr	Phe	Thr	Ser	Pro	Gly	Gly	Tyr	Tyr	Ala	Gly	Pro	Ala	Gly	Gly	85	90	95	
Asp	Pro	Gly	Ala	Phe	Leu	Ala	Met	Asp	Ala	His	Thr	Tyr	His	Pro	His	100	105	110	
Pro	His	Pro	Pro	Pro	Ala	Tyr	Phe	Gly	Leu	Pro	Gly	Leu	Phe	Gly	Pro	115	120	125	
Pro	Pro	Pro	Val	Pro	Pro	Tyr	Tyr	Gly	Ser	His	Leu	Arg	Ala	Asp	Tyr	130	135	140	
Val	Pro	Ala	Pro	Ser	Arg	Ser	Asn	Lys	Arg	Lys	Arg	Asp	Pro	Glu	Glu	145	150	155	160
Asp	Glu	Glu	Gly	Gly	Gly	Leu	Phe	Pro	Gly	Glu	Asp	Ala	Thr	Leu	Tyr				

100

Year	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Ser Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser Ser Ile
 5 10 15

Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala Thr Ala Ala Ala
 20 25

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Gly Val Pro Arg Arg Gln Arg Ala Ile Asp Lys Arg Gln Arg Ala
 5 10 15

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln
 5 10 15

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser
 5 10

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Pro Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala
 5 10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala
5 10

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser
5 10

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala
5 10

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala Thr
5 10

(2) INFORMATION FOR SEQ ID NO: 15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15

Arg Ala Ala Thr Ser Gly Ala Thr Ala Ala Ala Ser
5 10

(2) INFORMATION FOR SEQ ID NO: 16

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Ala Ala Val Asp Thr Gly Ser Gly Gly Gly Gly Gln
5 10

(2) INFORMATION FOR SEQ ID NO: 17

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Ala Val Asp Thr Gly Ser Gly Gly Gly Gln Pro
5 10

(2) INFORMATION FOR SEQ ID NO: 18

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Val Asp Thr Gly Ser Gly Gly Gly Gln Pro His
5 10

(2) INFORMATION FOR SEQ ID NO: 19

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Asp Thr Gly Ser Gly Gly Gly Gln Pro His Asp
5 10

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Gly Gly Gly Gln Pro His Asp Thr Ala Pro Arg Gly
 5 10

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg
 5 10

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys
 5 10